

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表平10-508967

(43) 公表日 平成10年(1998) 9月2日

(51) Int.Cl.⁸

G 0 6 F 12/00

識別記号

5 3 1

F I

G 0 6 F 12/00

5 3 1 D

審査請求 未請求 予備審査請求 未請求(全 34 頁)

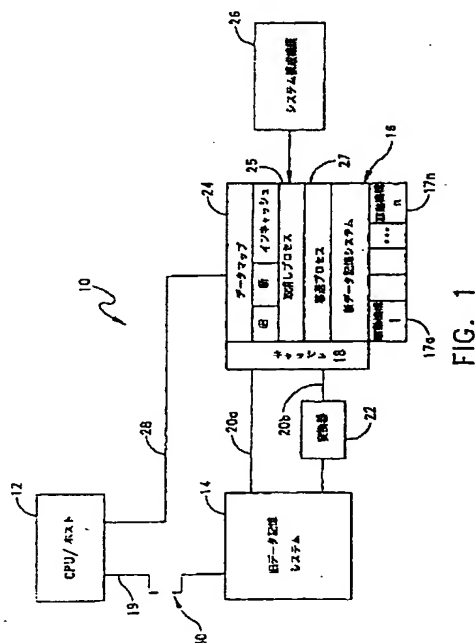
(21) 出願番号 特願平8-535206
 (86) (22) 出願日 平成8年(1996) 8月29日
 (85) 翻訳文提出日 平成9年(1997) 5月1日
 (86) 国際出願番号 PCT/US96/13781
 (87) 国際公開番号 WO97/09676
 (87) 国際公開日 平成9年(1997) 3月13日
 (31) 優先権主張番号 08/522, 903
 (32) 優先日 1995年9月1日
 (33) 優先権主張国 米国 (US)
 (81) 指定国 EP(AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, L U, MC, NL, PT, SE), JP, KR

(71) 出願人 イーエムシー コーポレイション
 アメリカ合衆国 マサチューセッツ州
 01748-9103 ホプキントン サウス ス
 トリート 171
 (72) 発明者 オフェク ユーヴァル
 アメリカ合衆国 マサチューセッツ州
 01748 ホプキントン フォレスト レー
 ン 13
 (72) 発明者 ヤナイ モーシェ
 アメリカ合衆国 マサチューセッツ州
 02146 ブルックリン カトリン ロード
 15
 (74) 代理人 弁理士 中村 稔 (外6名)

(54) 【発明の名称】 オンライン、実時間、データ移送の技術分野のためのシステム及び方法

(57) 【要約】

第1のデータ記憶システム(14)から第2のデータ記憶システム(16)までオンライン実時間透過データ移送を提供するためのシステム及び方法(25、27)。第2のデータ記憶システムは、ホスト(12)と第1のデータ記憶システム(14)の間に介在させられている。データマップ(24)が、第2のデータ記憶システム上に記憶されたデータ要素及び第1のデータ記憶システムから第2のデータ記憶システムまでコピーされた対応するデータ要素を識別する。ホストデータ要求に応じて、第2のデータ記憶装置は、中に記憶された場合のデータを検索する。そうでなければ、第2のデータ記憶システムは第1のデータ記憶システムからデータを検索し、データを自らに書き込み、データマップを更新する。要求を処理するに忙しくない場合、第2のデータ記憶システムは、結合されたあらゆるホストとは独立して第1のデータ記憶装置から第2のデータ記憶装置へとデータをコピーする。



【特許請求の範囲】

1. 第1及び第2のデータ記憶システムの間でオンライン透過データ移送を提供するシステムにおいて、
 - － 複数のデータ要素を保持する第1のデータ記憶装置；及び
 - － 前記第1のデータ記憶装置から第2のデータ記憶装置まで独立してデータを移送するように前記第1のデータ記憶装置に結合された第2のデータ記憶装置、を含んで成るシステム。
2. 前記第2のデータ記憶装置が外部のデータ発生源に結合されており、前記第1のデータ記憶装置から第2のデータ記憶装置までデータを移送するため前記外部データ発生源に対する応答性をもつ請求の範囲第1項に記載のシステム。
3. 前記データ記憶装置には、少なくとも1つの予め定められたデータ要素が前記第2のデータ記憶装置上に記憶されているか否かを表示するためのデータ要素マップが含まれている、請求の範囲第1項に記載のシステム。
4. 前記第2のデータ記憶装置は、前記外部発生源とは独立して前記第1のデータ記憶装置から前記少なくとも1つの予め定められたデータ要素を選択的に得るため、前記少なくとも1つの予め定められたデータ要素が前記第2のデータ記憶装置上に記憶されているか否かを表示する前記データ要素マップからの前記表示に対し応答性を有している、請求の範囲第3項に記載のシステム。
5. 前記第1のデータ記憶装置から前記少なくとも1つの予め定められたデータ要素を選択的に得るため及びこうして得られたデータを前記第2のデータ記憶装置上に記憶するため、前記第2のデータ記憶装置は、前記外部発生源によって発出された前記少なくとも1つの予め定められたデータをアクセスする読取り又は書込み指令の1つに対し応答性を有している、請求の範囲第4項に記載のシステム。
6. 第1及び第2のデータ記憶システムの間でオンライン透過データ移送を提供するシステムにおいて、
 - － 少なくとも1つのデータ記憶装置からデータを読み取りここへデータを書込むためのデータ処理装置；

- － 複数のデータ要素を保持する第1のデータ記憶装置；及び
- － 前記データ処理装置及び前記第1のデータ記憶装置に結合されている第2のデータ記憶装置、

を含んで成り、前記第2の記憶装置には、前記少なくとも1つの予め定められたデータ要素が前記第2のデータ記憶装置上に記憶されているか否かを表示するためのデータ要素マップが含まれており、前記第2の記憶装置は、前記少なくとも1つの予め定められたデータ要素が前記第2のデータ記憶装置内に記憶されているか否かの前記表示に対し少なくとも応答性を有しかつ前記第1のデータ記憶装置からデータを選択的に得るため及び前記第2のデータ記憶装置上に前記得られたデータを記憶するため前記データ処理装置によって発出された前記特定のデータ要素をアクセスする読取り又は書込み指令の1つに対し応答性を有している、システム。

7. 前記第2のデータ記憶装置は、前記データ処理装置から受理した予め定められたデータ要素についての読取り指令に対し応答性をもちかつ、少なくとも前記予め定められたデータ要素について前記第1のデータ記憶装置に対しデータ読取り指令を発出するため第2のデータ記憶装置上に前記予め定められたデータ要素が記憶されていないことの表示に対し応答性を有する、請求の範囲第6項に記載のシステム。

8. 前記第2のデータ記憶装置が前記第1のデータ記憶装置に対して複数の予め定められた要素についてのデータ読取り指令を発出する、請求の範囲第7項に記載のシステム。

9. 前記複数のデータ要素が複数の順次データ要素を内含している請求の範囲第8項に記載のシステム。

10. 前記第2のデータ記憶システムは、前記データ処理装置からより長期のデータ記憶装置へと予め定められた1つのデータ要素を書込むよう前記第2のデータ記憶装置を導く前記データ処理装置から受理されたデータ書込み指令に対して応答性をもち、しかも、少なくとも前記予め定められたデータ要素について前記第1のデータ記憶装置に対してデータ読取り指令を発出するため前記予め定められたデータ要素が前記第2のデータ記憶装置上に記憶されていないこと

の表示に対して応答性を有している、請求の範囲第6項に記載のシステム。

11. 前記第2のデータ記憶装置は、前記第1のデータ記憶システムに対し前記データ読取り指令を発出するため、前記データ処理装置から受理した前記データ書込み指令が、1つの完全データ要素記憶場所の一部分である予め定められたデータ要素についてのものであることの表示に対し応答性を有する、請求の範囲第10項に記載のシステム。

12. 前記第2のデータ記憶装置は、前記第2のデータ記憶装置に対して前記受理したデータ要素を書込むため、前記データ処理装置から受理した前記データ書込み指令が1つの全体的データ要素記憶場所である予め定められたデータ要素についてのものであることの表示に対し応答性を有する、請求の範囲第11項に記載のシステム。

13. 前記全体的データ要素記憶場所には、ディスク駆動機構トラックが含まれている、請求の範囲第12項に記載のシステム。

14. 前記第2のデータ記憶装置及び前記第2のデータ記憶装置データ要素マップに対してデータ記憶装置構成データを提供するため、前記第2のデータ記憶装置に結合されたデータ記憶装置構成装置をさらに含んでなる、請求の範囲第6項に記載のシステム。

15. 前記データ処理装置には少なくとも1つのホストコンピュータが含まれている、請求の範囲第6項に記載のシステム。

16. 前記データ処理装置には1つのネットワークが含まれている、請求の範囲第6項に記載のシステム。

17. 第1のデータ記憶装置から、1つの外部データ発生源及びこの第1のデータ記憶装置に結合された第2のデータ記憶装置までデータを移送するための方法において、前記第1のデータ記憶装置が複数のデータ要素を保持し、この複数のデータ要素のうちの少なくともいくつかが前記第2のデータ記憶装置にコピーされるべきものであり、

- 前記第1のデータ記憶装置上の少なくとも一定数のデータ要素記憶場所アドレスに対応する複数のデータ要素記憶場所アドレスを内含するべく前記第2のデータ記憶装置を構成する段階；及び

- ー 前記第 2 のデータ記憶装置により前記第 1 のデータ記憶装置から前記第 2 のデータ記憶装置へと少なくとも 1 つの予め定められたデータ要素を選択的にかつ独立してコピーする段階；

を含んで成る方法。

18. 前記第 2 のデータ記憶装置が、前記第 1 のデータ記憶装置から前記第 2 のデータ記憶装置へと少なくとも 1 つの予め定められたデータ要素を選択的にコピーするため、前記外部発生源に対し応答性を有する、請求の範囲第 17 項に記載の方法。

19. 前記外部発生源には 1 つのネットワークが含まれている請求の範囲第 17 項に記載の方法。

20. 前記第 2 のデータ記憶装置は、前記ネットワークとは独立して、前記第 1 のデータ記憶装置から前記第 2 のデータ記憶装置へと前記少なくとも 1 つの予め定められたデータ要素を選択的にコピーする、請求の範囲第 19 項に記載の方法。

21. 前記外部発生源には少なくとも 1 つのデータ処理装置が含まれている、請求の範囲第 17 項に記載の方法。

22. 前記第 2 のデータ記憶装置には、予め定められたデータ要素記憶場所アドレスをもつデータ要素が前記第 2 のデータ記憶装置上に記憶されているか否かの表示を少なくとも提供するデータ要素マップが含まれており、さらに

- ー 前記第 2 のデータ記憶装置により前記データ処理装置から、前記第 1 のデータ記憶装置上の予め定められたデータ要素記憶場所アドレスに記憶された少なくとも 1 つのデータ要素に関するデータ要素読取り要求及びデータ要素書込み要求の少なくとも一方を受理する段階；

- ー 前記少なくとも 1 つのデータ要素が前記第 2 のデータ記憶装置上に記憶されているか否かを決定するべく前記第 2 のデータ記憶装置により前記データ要素マップを探索する段階；及び

- ー 前記探索に応答して、前記第 1 のデータ記憶装置から前記第 2 のデータ記憶装置へ前記少なくとも 1 つのデータ要素を選択的にコピーする段階、

を含んで成る、請求の範囲第 21 項に記載の方法。

23. 予めデータ処理装置に結合された第1のデータ記憶装置から、前記データ処理装置に現在結合されている第2のデータ記憶装置及び前記第1のデータ記憶装置までデータを移送するための方法において、第1のデータ記憶装置には、前記データ処理装置によって以前にアクセスされた複数のデータ処理要素が含まれており、これらの複数のデータ要素の少なくともいくつかが前記第2のデータ記憶装置にコピーされるべきものであり、前記第2のデータ記憶装置には、予め定められたデータ要素記憶場所アドレスをもつデータ要素が前記第2のデータ記憶装置上に記憶されているか否かの表示を少なくとも含むデータ要素マップが含まれており、

- － データ要素記憶場所に対応する複数のデータ要素記憶場所アドレスを内含するべく前記第2のデータ記憶装置を構成する段階；
 - － 前記第2のデータ記憶装置により前記データ処理装置から、少なくとも1つのデータ要素に関するデータ要素読取り要求及びデータ要素書込み要求のうちの少なくとも1つを受理する段階；
 - － 前記少なくとも1つのデータ要素が前記第2のデータ記憶装置上に記憶されているか否かを決定するべく前記第2のデータ記憶装置により前記データ要素マップを探索する段階；及び
 - － 前記探索に応答して、前記第1のデータ記憶装置から前記第2のデータ記憶装置へ前記少なくとも1つのデータ要素を選択的にコピーする段階、
- を含んで成る方法。

24. さらに、

- － 前記データ要素要求がデータ読取り要求であることを前記第2のデータ記憶装置によって決定し、
 - ・ 前記要求された少なくとも1つのデータ要素が前記第2のデータ記憶装置上に記憶されているか否かを決定する段階、
 - ・ 前記要求された少なくとも1つのデータ要素が前記第2の記憶装置上に記憶されている場合、前記データ処理装置に対してこの要求された少なくとも1つのデータ要素を提供する段階
- を実行する段階、

- ー 前記要求された少なくとも1つのデータ要素が前記第2のデータ記憶装置上に記憶されていない場合には、
 - ・ 前記データ処理装置によって要求された前記少なくとも1つのデータ要素について前記第1のデータ記憶装置に対しデータ要素要求を生成する段階；
 - ・ 前記第1のデータ記憶装置により前記少なくとも1つのデータ要素を検索し、前記第2のデータ記憶装置に対し前記検索された少なくとも1つのデータ要素を提供する段階；を実行する段階、
- を含んで成り、

前記第2のデータ記憶装置は前記第1のデータ記憶装置から前記検索された少なくとも1つのデータ要素を受理し、前記要求された少なくとも1つのデータ要素を前記データ処理装置に提供する、

請求の範囲第23項に記載の方法。

25. さらに、

- ー 前記第2のデータ記憶装置により、前記データ要素要求がデータ書込み要求であるか否かを決定し、
 - ・ 書込むべき前記少なくとも1つのデータ要素が現在前記第1の又は前記第2のデータ記憶装置上に記憶されているか否かを決定する段階、
 - ・ 前記要求された少なくとも1つのデータ要素が前記第2のデータ記憶装置上に記憶されている場合、前記少なくとも1つのデータ要素を前記第2のデータ記憶装置に書込む段階を実行する段階、
- ー 前記少なくとも1つの書込むべきデータ要素が前記第2のデータ記憶装置上に記憶されていない場合、
 - ・ 前記データ処理装置により書込まれるよう要求された前記少なくとも1つのデータ要素について前記第1のデータ記憶装置に対してデータ要素読取り要求を生成する段階；
 - ・ 前記第1のデータ記憶装置により前記少なくとも1つのデータ要素を検索

し、この検索された少なくとも1つのデータ要素を前記第2のデータ記憶

装置に提供する段階

を実行する段階、

を含んで成り、

前記第2のデータ記憶装置は前記受理された少なくとも1つのデータ要素を前記第1のデータ記憶装置から受理し、

前記第2のデータ記憶装置は、書込まれるべき前記少なくとも1つのデータ要素を前記第2のデータ記憶装置に書込む、

請求の範囲第23項に記載の方法。

26. 前記第1のデータ記憶装置が前記データ処理装置に対し予め結合されており、前記第2のデータ記憶装置を前記データ処理装置に結合する前に前記データ処理装置から前記第1のデータ記憶装置を結合解除する段階を実行する、請求の範囲第23項に記載の方法。

27. 前記第2のデータ記憶装置により前記第1のデータ記憶装置内に記憶された前記要求された少なくとも1つのデータ要素を受理する前記段階には、キャッシュメモリ内に前記受理された要求された少なくとも1つのデータ要素を記憶することが含まれる、請求の範囲第23項に記載の方法。

28. 前記第1のデータ記憶装置から前記要求された少なくとも1つのデータ要素を受理する前記段階の後、前記データ要素マップを更新する段階をさらに含んで成る、請求の範囲第23項に記載の方法。

29. 前記データ処理装置に対して前記要求された少なくとも1つのデータ要素を提供する段階の後、前記要求された少なくとも1つのデータ要素を前記第2のデータ記憶装置に書込む段階をさらに含んで成る、請求の範囲第23項に記載の方法。

30. 一少なくとも前記データ処理装置からのデータ要素読取り及び書込み要求に応えるのに前記第2のデータ記憶装置が完全に忙しい状態にないことを決定する段階；

一 前記第2のデータ記憶装置の前記データ要素マップを読取る段階；

- ー 前記第1のデータ記憶装置上に記憶されたどのデータ要素が前記第2のデータ記憶装置にコピーされていないかを決定する段階；
 - ー 前記第1のデータ記憶装置から前記第2のデータ記憶装置へコピーされなかった前記データ要素を要求する段階；
 - ー 前記第2のデータ記憶装置上に前記要求されたデータ要素を記憶する段階；及び
 - ー 前記コピーされたデータ要素が前記第2のデータ記憶装置上に記憶されていることを表示するべく前記データ要素を更新する段階、
をさらに含んで成る請求の範囲第23項に記載の方法。
31. 前記第2のデータ記憶装置のチャネルプロセスは、一連のデータ要素に対する一連の要求が順次データ要素要求となるという表示を前記第1のデータ記憶装置に対して提供する、請求の範囲第23項に記載の方法。
32. 前記データ処理装置により要求された前記少なくとも1つのデータ要素について前記第1のデータ記憶装置に対してデータ要素要求を生成する段階には、前記第1のデータ記憶装置により理解された1つのフォーマットで前記データ要素要求を生成することが含まれている、請求の範囲第24項に記載の方法。
33. データ処理装置に予め結合された第1のデータ記憶装置から第2のデータ記憶装置までデータを移送するための方法において、第1のデータ記憶装置には、前記データ処理装置によって現在アクセスされつつある複数のデータ処理要素が含まれ、そこからこれらの複数のデータ要素の少なくともいくつかのコピーされるべきものであり、
- ー 前記データ記憶装置によってアクセスされるべき前記複数のデータ要素を記憶するため、1つのデータ要素が上に記憶されているか否かの表示を少なくとも1つ含むデータ要素マップを内含する第2のデータ記憶装置を提供する段階；
 - ー 前記データ処理装置に対して前記第2のデータ記憶装置を結合する段階；
 - ー 前記第2のデータ記憶装置を前記第1のデータ記憶装置に結合する段階；
 - ー 前記第1のデータ記憶装置上のデータ要素記憶場所に対応する複数のデー

タ要素記憶場所アドレスを内含するように前記第2のデータ記憶装置を構成する段階；

－ 前記第2のデータ記憶装置が前記データ処理装置からのデータ要素要求に

応えるのに忙しくないことを決定する段階；

－ 前記第2のデータ記憶装置の前記データ要素マップを読取る段階；

－ 前記第1のデータ記憶装置上に記憶されたデータ要素が前記第2のデータ記憶装置上に記憶されていないことを決定する段階；

－ 前記第1のデータ記憶装置から前記第2のデータ記憶装置にコピーされなかった前記データ要素を要求する段階；

－ 前記第2のデータ記憶装置上に前記要求されたデータ要素を記憶する段階；及び

－ 前記記憶されたデータ要素が前記第2のデータ記憶装置上に記憶されていることを表示するべく前記データ要素マップを更新する段階、

を含んで成る方法。

【発明の詳細な説明】

オンライン、実時間、データ移送の技術分野のためのシステム及び方法

本発明は、データ記憶システム、より特定的には、既存のデータ記憶サブシステムのオンライン置換のためのシステム及び方法に関する。

背景技術

例えば銀行、航空会社及び保険会社といったような事業及び組織のデータ処理センターは、データ記憶装置上に記憶された大量のデータにアクセスし処理する自らの能力にほぼ排他的に依存している。より大型のデータ記憶システムの一部をなす単数又は複数のデータ記憶装置上に標準的に記憶されているデータ及びその他の情報は、一般にデータベースと呼ばれる。

データベースはほぼ常に「開放した」状態にあり、恒常的に「使用中」であり、結合されたデータベース処理システム、中央処理ユニット（CPU）又はホスト本体コンピュータによってアクセスされている。データアクセス不能は、かかる事業及び組織にとっては危機的とまでは言わなくとも、莫大な損害をもたらすことであり、典型的には、事業又は組織が一時的に操業を停止せざるを得なくなるという結果をもたらすことになる。

通常の操業中、これらの事業又は組織は、そのデータ記憶装置及びデータ記憶システムをグレードアップしなければならない。このようなグレードアップは時として、既存の物理的システムにデータ記憶容量を付加することしか含んでいないが、通例では、グレードアップによって完全に別の新しいデータ記憶システムの付加が求められることになる。このような場合、既存のデータ記憶システム又は装置に対する既存のデータは、テープ駆動機構といった別の装置上でバックアップされなくてはならず、新しいシステムはデータ処理ユニットに設置され接続され、又データはバックアップ装置から新しいデータ記憶システムへとコピーされなくてはならない。このような活動は物理的に、達成するのに少なくとも2日がかかる。転換に2日以上かかる場合又は事業又は組織が2日の操業不能に耐えることができない場合、そのデータ記憶システムをグレードアップする必要性及

び望みは、克服不可能な問題に直面する可能性がある。

互いに結合されるべき第1のシステム及び第2のシステムという同じタイプの2つのデータ記憶システムを許容し、ポストデータ処理システムからの介入又はこれとの干渉無しに第1のシステムから第2のシステムまでのデータのコピー作業をデータ記憶システム自体が制御できるようにするいくつかの先行技術に基づくデータコピー方法及びシステムが提案されてきた。例えば、EMC Corporation, Hopkinton, Massachusettsから入手可能なSymmetrix 5500データ記憶システム上で実現可能である1つのこのような遠隔データコピー設備の機構について記述する、本書にその全文が参考として内含されている「遠隔データミラリング」という題の米国特許出願第68/052,039号の中に記述されたデータ記憶システムを参照されたい。

データコピーのためのこのようなシステム及び方法が可能であるものの、大部分のケースにおいて、第1及び第2のデータ記憶システムは同じタイプのものではないか、或いはデータベースが開放している間にホストによる補助無しで2つのデータ記憶システムの間でこのような「背景」データ移送を行なうことを可能にするタイプのものではない。その上、このような先行技術のデータ記憶システムの上でさえ、データベースが「開放した」状態にある間に「背景」タスクとしてデータを移送することには、ホスト又は中央処理ユニットによってアクセスされるにつれてデータがつねに変更しており、従って、旧システムがホストに接続された状態に放置された場合には旧データ記憶システム上に記憶されたデータと新しいデータ記憶システム上に移送されたデータの間につねに不均衡が存在することになる、という事実が考慮に入れられていない。このような場合、新しいデータ記憶システムは決して完全に「追いつく」ことができず、旧データ記憶システムに対して完全に同期化され得ない。

従って、必要とされるのは、データベースが開放状態にある間に、実時間でかつホスト又はデータ処理ユニットに対して完全に透過性のデータ移送を第1のデータ記憶システムと第2のデータ記憶システムの間で可能にするためのシステム及び方法である。

発明の開示

本発明は、2つのデータ記憶装置の間でオンライン実時間透過データ移送を提供するためのシステム及び方法に関する。このシステムにはホストコンピュータといったようなデータ処理装置を含む外部のデータ発生源に予め結合された第1のデータ記憶装置、又は一定数のホストコンピュータといったような一定数のデータ処理装置に接続され得るネットワークが内含される。ホストコンピュータといったようなデータ処理装置は、データ記憶装置からデータを読みとり、この記憶装置へデータを書き込む。第1のデータ記憶装置は、最初、現在データ処理装置によってアクセスされている複数のデータ要素を内含する。

データ処理装置によってアクセスされるべきデータ要素を記憶するために、第1にデータ記憶装置及びデータ処理装置に結合されている少なくとも1つの第2のデータ記憶装置が具備されている。この第2のデータ記憶装置は、好ましくは、特定のデータ要素が第2のデータ記憶システム上に記憶されているか否かの表示を少なくとも1つ含むデータ要素マップを内含している。

一つの実施形態においては、第2のデータ記憶システムは、発生源とは独立して第1のデータ記憶システムから第2のデータ記憶システムまでデータを独立的に移送する。もう一つの実施形態では、第2のデータ記憶システムは、データを第1のデータ記憶システムから第2のシステムまで移送するため、外部の発生源に対する応答性をもつ。

さらにもう一つの実施形態においては、データ処理装置は、データ読取り要求（読取りデータオペレーションの場合）又はデータ書込み指令（書込みオペレーションの場合）を発出する。この要求は、第2のデータ記憶装置によって受理される。読取りオペレーションの場合、第2のデータ記憶装置は、データが第2のデータ記憶装置まで移送されてこの上に記憶されているか否かを決定するべくデータマップ又はテーブルを検査する。データが第2のデータ記憶装置上に記憶されていることが決定されたならば、そのデータは要求している装置が利用できる状態となる。

データが第2のデータ記憶装置上に記憶されていない場合、第2のデータ記憶装置は、読取りデータ指令の形で第1のデータ記憶装置に対してデータ要求を発

出し、データを獲得し、要求している装置がそのデータを利用できる状態にする。第1のデータ記憶装置から受理したデータも同様に第2のデータ記憶装置に書込まれ、データマップは更新される。

書込みオペレーションの場合、1つの実施形態では、データ処理装置から受理されたデータが古い方つまり第1のデータ記憶装置（データマップ内に「移送要」とマークされたデータ記憶場所）からまだコピー又は「移送」されていないデータ記憶システム上の1つの場所を宛先としているものであり、データが全又は完全データ要素でない（例えばデータの「全トラック」でない）場合、書込みオペレーションが中断され、第1のデータ記憶装置上の対応する場所（例えば「全トラック」）からの「完全な」データ要素は第2のデータ記憶装置上のキャッシュメモリ内に読み取られ、インーキャッシュフラグ又はビットがセットされ、データ記憶場所が「書込み保留」としてマーキング又は識別され、書込みオペレーションが再開されて、データが第2のデータ記憶システムのキャッシュメモリ内に現在記憶されているデータの「全トラック」に対し又はこの全体にわたって書込まれることになるということを意味するように考慮されている。その他の実施形態では、書込むべき新しいデータが完全データ要素（例えば「全トラック」）であることがわかっている場合、第1のつまり古い方のデータ処理装置から、古い方のデータを検索することはできない。

第2のデータ記憶装置が、ホストコンピュータといったような結合されたデータ処理装置からのデータ読取り又は書込み要求を処理するのに忙しくない場合、第2のデータ記憶システムは、どのデータ要素が第1のデータ記憶装置上で常駐であり第2のデータ記憶装置上には記憶されていないかを決定するためそのデータマップ／テーブルを検査する。次に、第2のデータ記憶装置は、これらのデータ要素のうちの1つ又は複数を要求して第1のデータ記憶装置に対し読取り要求を発出し、データを受理し、データを第2のデータ記憶装置に書込み、そのデータが今第2のデータ記憶装置上に記憶されていることを表示するためデータマップ／テーブルを更新する。

このようにして、第1及び第2のデータ記憶装置の間で時間のかかるオフラインデータ移送を行なう必要は全く無く、むしろデータのコピー又は移送は、デー

タ記憶装置がオンラインでホスト又はその他の要求している装置が利用できる状態にある間に、実時間でかつ結合されたデータ記憶装置に対して完全に透過性ある形で行なうことができる。

好ましい実施形態においては、第2のデータ記憶装置はさらに、第2のデータ記憶装置のデータ要素マップ又はテーブルに対して構成データを提供し、第2のデータ記憶装置が一般に第1のデータ記憶装置と類似した又は同一であるような要領で少なくとも部分的には構成され得るようにするコンピュータといったようなデータ記憶装置システム構成装置を内含するか、又はこれに結合されている。

付加的には、好ましい実施形態では、第2及び第1のデータ記憶装置が「ESCON」通信プロトコルを利用する光ファイバリンクといったような高速通信リンクによって結合されているように考慮されている。好ましい実施形態では同様に、データ記憶装置がディスク装置といったような複数のデータ記憶装置を内含するように考えられている。この場合、データ要素は、単数又は複数のディスク駆動機構ボリューム、トラック又はレコードを内含できる。

図面の簡単な説明

本発明のこれらの及びその他の特徴及び利点は、図面と共に以下の詳細な記述を読むことによってより良く理解できることだろう。なお図中

図1は、本発明に従った第1及び第2のデータ記憶システムの間のオンライン、データ透過性データ移送を提供するためのシステム及び方法を達成することのできるデータ処理及びデータ記憶システムの一例の概略図である。

図2は、データ要素マップ又はテーブルの概略図である。

図3は、本発明の方法に従った第1及び第2のデータ記憶システムの間のオンライン透過データ移送を提供する各段階を概略的に示す流れ図である。

図4は、第2のデータ記憶装置がホスト又はデータ処理装置からのデータ要求を取扱うのに忙しくない時に、データ記憶装置又はホストシステムの介入無く第1及び第2のデータ記憶システムの間でのデータ移送を提供するための各段階を例示する流れ図である。

本発明の最良の実施形態

本発明は、ホストコンピュータなどのデータ処理装置に結合されているものを少なくとも1つ含む2つのデータ記憶システムの間のオンライン、実時間透過データ移送を提供するためのシステム及び方法に関する。

本発明を実現、実施することのできる図1のシステム10の例には、ホストコンピュータ、中央処理ユニット又はその他の類似のデータ処理装置12が含まれている。データ処理装置12は最初第1のデータ記憶システム14に結合されている。ほとんどの場合において、第1のデータ記憶システム14は、データ処理装置12のニーズを取扱うのに十分な大きさをもたないか又は他の何らかの理由で完全に又は部分的に置換されるか又は第2のデータ記憶システム16の付加により増大されようとしている、古い方のデータ記憶システムである。

第1のデータ記憶システム14は当初、データ通信リンク19を用いてデータ処理装置12に結合されている。第2のデータ記憶システム16は、単数又は複数のデータ通信経路20a及び／又は20bを用いて第1のデータ記憶システム14に結合されている。データ通信経路20a-20bの例は、当業者にとって周知のIBM「母線及びタグ」接続及びESCONデータ接続といったより高速の光ファイバ接続を内含する。

第1及び第2のデータ記憶システム14、16が互換性の無いデータ通信プロトコル又はインターフェイスを有する場合、必要に応じて、当該技術分野において周知の通りに、単数又は複数のデータ通信リンク20a、20b上にプロトコル変換器22を具備することができる。

第2のデータ記憶システム16には、少なくとも第2のデータ記憶システム16上に記憶されているデータ要素のデータマップ又はテーブル24が含まれている。このデータマップ又はテーブルは、第2のデータ記憶システム16のセットアップ又は構成中に作成され、第2のデータ記憶システム16の特定の構成により左右される。

好ましくは、データマップ／テーブル24は同様に、第1のデータ記憶システム14の中に記憶されているデータ要素についての情報も内含しており、かかるデータマップ／テーブルの使用について以下でさらに詳細に説明する。

第2のデータ記憶システム16は標準的にかつ好ましくは、ユーザーが自らの望むとおりに第2のデータ記憶システム16及びデータマップ／テーブル24を構成することができるようにするコンピュータといったようなデータ記憶システム構成装置26に結合されている。好ましい実施形態においては、第2のデータ記憶システム16は、少なくとも部分的に、論理装置数、記憶サイズ、記憶システムタイプ（例えば3380／3390）などに関して第1のデータ記憶システム14が構成されているのと全く同じ様に構成される。

好ましい実施形態においては、データ記憶システム構成装置26は、ユーザーが、第1のデータ記憶システム14上のデータ要素記憶アドレスに対応するデータ要素記憶場所又はアドレスを内含するように第2のデータ記憶システム16上にデータ記憶領域の少なくとも一部分を構成することができるようにする。

好ましい実施形態においては、第2のデータ記憶システム16は、固定ブロックアーキテクチャ（FBA）でフォーマットされたディスク駆動機構17a-17nを多数利用し、ホストコンピュータ又はその他のデータ処理装置12によってアクセスされるべき大量のデータを記憶するように適合されたディスク駆動機構データ記憶システムである。第2のデータ記憶システム16の例には、第2のデータ記憶システム16とホスト又はその他のデータ処理装置12の間でデータ読取り及び書込み要求を保持又は緩衝するのに役立つキャッシュメモリー18も標準的に含まれている。このようなデータ記憶システムは、当業者には周知のものであり、例えば、本書にその明細書が参考として内含されているEMC Corporation, Hopkinton, Massachusettsから入手可能なSymmetrix 5500シリーズデータ記憶システムを含む。

当初、第2の、つまり新しいデータ記憶システム16は、単数又は複数のデータ通信リンク又は経路20a、20bを用いて第1のデータ記憶システム14にまず結合される。システム構成装置26またはその他の類似の又は等価の装置を用いて、又はホスト12によって第2のデータ記憶システム16が構成された後、この第2のデータ記憶システム16は、ホストコンピュータ12又はその他のデータ処理装置へ、データ通信経路28を用いて結合される。

好ましくは、データ通信経路28は、光ファイバ「ESCON」通信経路とい

ったような高速通信経路であるが、他のいかなる通信経路でも全て本発明の範囲内に入るものとみなされる。第2のデータ記憶システム16内でホスト又はその他のデータ処理ユニット12の間でデータ通信経路28を接続する直前に、ホスト12と第1のデータ記憶システム14の間の予め既存のデータ通信経路18は、矢印30で例示されている通り接続解除又は切断される。

かくして、第1のデータ記憶システム14上のデータをバックアップしその後ひき続き第1のデータ記憶システム14を第2のデータ記憶システム16で置換し、その後全てのデータを新しいデータ記憶システム16又はもとの「第1の」データ記憶システム14に結合された状態にとどまっているホストに対してコピーすることができるようにするため一定の日数だけホスト又はその他のデータ処理システム12をラインから外さなくてはならない先行技術とは異なり、本発明では、第1のデータ信号経路19が切断又は接続解除され第2のデータ信号経路28が第2の又は新しいデータ記憶システム16とホストコンピュータ又はその他のデータ処理装置12の間に樹立されている間、ホストコンピュータ又はその他のデータ処理装置12を比較的短時間（この手順には標準的に約10分以下の時間が必要である）ラインから外し又サービスを中断させることしか必要とされない。

従って、第2のデータ記憶システム16がホスト又はその他のデータ処理ユニット12に接続された後、ホスト又はデータ処理ユニット12が「その」データ記憶システムからデータを読み取るか又はこのシステムにデータを書込む要求を発出した時点でつねに、この要求は第2のデータ記憶システム16によって受理される。予め作成され構成されたデータマップ／テーブル24からのビット又はフラグを用いて、第2のデータ記憶システム16は、データマップ／テーブル24を走査することにより、要求されたデータ（読取りオペレーションの場合）が第1のデータ記憶システム14上又は第2のデータ記憶システム16上に記憶されるか否かを決定する。

このような階層データマップ／テーブル24については、さらに本書中ならびに、本発明の譲受人に対し譲渡され本書に共に参考として全体的に内含されている米国特許第5,206,939号及び第5,381,539号の中に説明され例示

され

ている。

データがすでに第2のデータ記憶システム16の中に記憶されている場合、第2のデータ記憶装置16は、当該技術分野において周知の通りにデータを検索し（恐らくはデータをキャッシュメモリー18内に一時的に記憶させる）、ホスト又はその他の要求しているデータ処理装置12がデータをいつでも利用できる状態にする。

要求されたデータが第2のデータ記憶システム16上に無い場合、第2のデータ記憶システム16のチャンネル又は実時間データ処理プロセス25が、第1のデータ記憶システム14（例えば標準IBMデータ読取り指令）に固有の又は既知の要領及びフォーマットで第1のデータ記憶システム14に対して読取りデータ要求を発出する。チャンネル又は実時間データ処理プロセス25は、好ましい実施形態においては、ホスト又はCPUに対する第2のデータ記憶システムのインターフェイス（標準的に「チャンネル」と呼ばれる）から単数又は複数の指令を受理し、これらの指令を解釈し、第1のデータ記憶システムが作用を及ぼし得る単数又は複数の対応する指令を発出する。このような「インタプリタ」タイプのソフトウェアは、当業者にとって周知のものである。

このとき、第1のデータ記憶システム14は要求されたデータを検索し、それを第2のデータ記憶システム16に提供する。第2のデータ記憶システム16は次に、そのデータを要求したホスト又はその他のデータ処理ユニット12がこのデータを利用できるようにする。

ここで第2のデータ記憶システムはデータのコピーを有することから、このデータは第2のデータ記憶システム16に書込まれることになり、適切なデータマップ/テーブル24のフラグ又はビットは更新されて、データが第2のデータ記憶システムまで移送されており従って次に同じデータ要素が要求されたときには第2のデータ記憶システム16がシステム上にすでに記憶されたデータを有しそれを第1のデータ記憶システムから要求しなくてすむということを表示することになる。

さらに、以下でより詳しく説明する通り、第2のデータ記憶システム16は「背景」データ移送手順又はプロセス27を実行することができる。プロセス

27の「背景」データ移送手順は、好ましい実施形態においては、データ移送を調整、監視及び制御し、かくして第2のデータ記憶システムがホスト又はその他のデータ処理装置12からのデータ入力／出力要求を取扱うのに忙しくない場合にはつねに、第2のデータ記憶システム16の移送プロセス27がそのデータテーブル／マップ24上の特定のフラグ又はビットを読取ることによって第1のデータ記憶システム上のどのデータがコピーされていないかを決定し、第1のデータ記憶システム14から第2のデータ記憶システム16までホスト12に対し完全に透過的にしかも往々にしてホスト又はその他のデータ処理装置12によるデータに対する完全なアクセス可能性を維持しながらホスト又はCPU12からのデータを検索しつつある可能性のあるチャネルプロセス25と並行してデータをコピー又は「移送」することになる一連の命令を含むソフトウェアプログラムである。

データ要素マップ／テーブル24の一例は、図2にさらに詳細に示されている。好ましい実施形態においては、データマップ／テーブル24は階層的に組織される。例えば、データ記憶システムがディスク駆動機構17a-17nといった複数のより長期のデータ記憶装置を含み、各々のディスク駆動機構が単数又は複数の論理「ボリューム」に分割され各ボリュームが一定数のディスク駆動機構トラックを内含している好ましい実施形態については、データマップ／テーブル24は、まず第1にディスク駆動機構といった各々の物理及び／又は論理装置のためのエントリ50を有することになる。

装置エントリ50の後には、第1の論理ボリュームのためのエントリ52が続く、その後には、論理ボリューム52を含む装置の各トラックのための単数又は複数のエントリ54a-54cが続く。第1の論理のためのエントリ52、54a-54cの後には、ライン50でのエントリにより表示された物理装置上で構成された第2の論理ボリュームのためのエントリライン56が続くことになる。

「データ・イン・キャッシュ」表示フラグ又はビット58を除くデータ記憶シ

システム内の各装置及びデータ記憶システムについての全ての情報は、データマップ／テーブル24内に階層フォーマットで記憶されている。従って、第2のデー

タ記憶システム16が特定のデータ要素についての情報（それが個々のデータレコード、トラック又はボリュームであるか否かに関わらず）を得たい又は得ることを必要としている場合にはつねに、データ記憶システム16は、望まれる基準又は特徴が装置のいずれかのトラック又はボリュームについて設定されたか否かを決定するため、装置レベル50で始めてデータマップ／テーブル24を走査する。

このデータマップ／テーブル24によって表わされたデータ記憶装置の一部分の中に望まれる特徴が見い出される場合、装置エントリ50、ボリュームエントリ52及び該当するトラックエントリ54内に、「フラグ」又はその他の類似した標識ビットセット又はその他の指示が存在することになる。

例えば、データマップ／テーブル24の好ましい実施形態には、第2のデータ記憶システム16のキャッシュ18内に現在特定のデータ要素が記憶されておりディスク駆動機構17a-17nといったより長期の記憶装置に書込まれなくてはならない場合にセットされる書込み保留フラグ又はビット61が含まれている。一例として、ボリューム1のトラック2が第2のデータ記憶システム16内のキャッシュ18内にあり書込み保留中であると仮定すると、ライン50での装置の書込み保留ビット61と同様に、データマップ／テーブル24のライン52におけるボリューム1の書込み保留ビット61と同様に、ラインエントリ54bにおけるインキャッシュビット58（トラック2について）及び書込み保留フラグ2はビット61がセットされることになる。

かくして、要求された特定のトラック又はレコードが書込み保留中であるか又は第2のシステムまで移送されてしまったか又はその他の何らかの属性又は特徴をもつものであるか否かを第2のデータ記憶システム16が決定することを望む場合、このシステム16は、まず最初にデータ要素がどの装置又はディスク駆動機構17a-17nに記憶されているかを決定し、次にその装置についての該当する標識フラグビットを検査する。その装置について特定の標識フラグビットが

セットされない場合には、第2のデータ記憶システム16は直ちに、その装置内のボリューム又はトラックといったようなより低いレベルの記憶ユニット又は記憶場所のいずれもその属性をもたないということがわかる。トラック又はボリューム

といったような階層構造内のいずれかの下位のデータ記憶要素が属性を含む場合、装置についてのその属性又はフラグビットがセットされることになる。

同様にして、論理ボリュームの一部であるレコード又はトラックといったような特定のデータ記憶場所が要求された属性を有する場合、そのボリュームに対する対応する属性又はフラグビットがセットされることになる。かくしてデータ記憶システム16は迅速に、検査中のボリューム又はその他の類似の論理又は物理区分よりも低いレベルをもついずれかのデータ記憶場所が特定の属性を有するか否かを、各々及び全ての下位レベルのデータ記憶場所を走査又は検索することなく決定することができる。

「インキャッシュ」フラグ又はビットは、データマップ／テーブル24の各々のライン又はエントリ50-56が直接アドレス可能であり、特定のデータ要素が現在「インキャッシュ」であるか否かを問い合わせるか又は「ルックアップ」しなければならない場合に第2のデータ記憶システムがその特定のデータ要素についてのテーブルエントリラインを直接アドレスする、という点において、階層構造に対する1つの例外である。しかしながら、このフラグ又はビットは本発明の範囲から逸脱することなく階層的な形で管理することもできるということが理解できる。

インキャッシュビット又はフラグ58及び書込み保留フラグ又はビット61に加えて、本発明の1つの特徴であるデータマップ／テーブル24は、好ましい実施形態において、無効トラックフォーマットフラグ又はビットといったようなその他のフラグビット62、及び、一般にフラグ又は標識ビット60によって示されるように、特定の装置、ボリューム又はトラック上のデータが移送を必要とするか、又はそれぞれ第1のデータ記憶システムから第2のデータ記憶システム14/16まですでに移送されたかの表示を含んでいる。

データマップ／テーブル24はさらに、新しい又は第2のデータ記憶システム

16のディスク駆動機構17a-17n上で対応するデータ要素を見い出すことのできる開始データアドレス64を識別するマップまたはテーブル24内の各々の要素についての物理アドレス64のエントリも含むことができる。

ここで、本発明に従った方法の作動について、ホスト又はその他の処理装置

12からのデータ要素読取り又は書込み要求を第2のデータ記憶システム16が受理する(ステップ100)図3のステップ100から始めて、詳細に説明する。この方法では次に、要求又は指令が読取り又は書込み要求のいずれであるかを決定する(ステップ101)。指令が読取り指令である場合、第2のデータ記憶システム16のチャンネル取扱いプロセス25は次に、要求されたデータが第2のデータ記憶システム16内にすでに記憶されているか否かを、そのデータテーブルマップ/テーブル24を読取ることによって決定する(ステップ102)。

第2のデータ記憶システム上にデータが記憶されている場合(ステップ102)、第2のデータ記憶システム16はホスト又はその他の要求中のデータ処理装置12がそのデータを利用できるようにし(ステップ104)、ステップ100に戻って新しいデータ読取り又は書込み要求の受理を行う。

しかしステップ102において第2のデータ記憶システム16が、その上にデータが現在記憶されていないということを決定した場合には、システム16は第1のデータ記憶システム14に対してデータを読取るよう要求を生成する(ステップ106)。

第1のデータ記憶システム14からデータを読取るための指令又は要求は、ホスト12から発出されるデータ読取り指令と同じ形をとる。かくして例えば、ホスト12がIBM又はIBMと互換性あるホスト又はデータ処理装置である場合、第2のデータ記憶システム16は第1のデータ記憶システム14に対しIBM互換性の「読取り」指令を発出することになる。第2のデータ記憶システム16のチャンネル及び移送プロセス25、27は、第1のデータ記憶システム14に固有の指令リストを維持し、必要とあらば、ホスト12により発出され第2のデータ処理システム16によって理解された第1の指令タイプから第1のデータ記憶システム14によって理解された第2の指令タイプへと指令タイプを容易に変換

することができる。

その後、第2のデータ記憶システム16は、第1のデータ記憶システム14から要求されたデータを受理し（ステップ108）、データ要素マップ／テーブル24を更新している間に第2のデータ記憶システム16のキャッシュメモリ18に対しデータを書込む（ステップ110）。このとき、第2のデータ記憶システ

ム16は、データがいつでも読取ることのできる状態にあることの表示を、ホスト又はデータ処理装置12に提供する。それに続いて、第2のデータ記憶システム16はキャッシュメモリ18からのデータを、第2のデータ記憶システム16上のディスク駆動機構といったより永久的な記憶場所へ書込み（ステップ114）、その後、データ要素マップ／テーブル24の単数又は複数のビット又はフラグへの最終的更新が行なわれる（ステップ116）。

従って、要求されたデータが第2のデータ記憶システム16上にまだ記憶されていない場合、ホスト12からの「読取り要求」指令の結果、第2のデータ記憶システム16は第1のデータ記憶システム14から第2のデータ記憶システム16までデータを「移送」することになる。

ホスト又はその他のデータ処理システム12が書込み要求又は指令を発出したならば（ステップ120）、第2のデータ記憶システム16のチャネルプロセス25は、書込むべきデータが第1のデータ記憶システムから第2のデータ記憶システムまで以前に移送されているか否かを決定する（ステップ122）。データが以前移送されていた場合（ステップ122）、第2のデータ記憶システムはデータをキャッシュに書込み、データマップ／テーブル24内のいずれかの必要なフラグ又はビットを更新する（ステップ110）。処理は、以前に記述した通りに続く。

しかしながら、データが以前に移送されていない場合（ステップ122）、本発明の方法は次に、ホストが発出する指令又は要求のタイプにより（例えばIBMホスト指令の場合）、書込み要求が全又は完全データ「トラック」といった全ての又は完全なデータ要素記憶場所についてのものであるか否かを決定する（ステップ124）。書込み要求が、全「トラック」又はその他の類似のタイプのデ

ータブロック又は内容についてのものである場合、第2のデータ記憶システムは、全ての「旧」データが現行の指令によって置換されつつあることから、データを第1のデータ記憶システム14から移送する心配をする必要はなく、従って処理は以前に記述した通りステップ110まで続く。

しかしながら、この方法により、書込み要求がトラックといった全又は完全データブロック又は領域より小さいものについてのものであるということが決定さ

れた場合（ステップ124）、方法は次に書込み要求の取扱いを一時的に中断して（ステップ126）、全又は完全「トラック」のための「読取り」指令を発出し、予め定められた量のデータ（例えばデータトラック全体）を読取り（ステップ128）、全データ「トラック」を第2のデータ記憶システム16のキャッシュメモリ18にコピーする。このとき、書込むべき新しいデータキャッシュメモリ18内の適切な記憶場所に書込まれ（実際の「書込み」指令の出現）、データテーブル/マップ24は（例えば、データがキャッシュメモリ18内にあること〔データインキャッシュビットセット〕、書込みがこのデータ上で保留されていること〔書込み保留ビットセット〕及びデータ要素が移動させられたこと〔データ移送要ビットリセット〕を表示するように）更新され、ホスト又はその他の中央処理ユニット12には書込み指令が完全であることが知らされる。

その後の或る時点で、書込み保留としてフラグ付けされたキャッシュメモリ18内のデータは、ディスク駆動機構といったより永久的な記憶場所にコピーされ、書込み保留ビットがリセットされる。

標準的には、データ書込み要求は、予め定められたデータ記憶要素又は物理/論理領域（例えばディスク駆動機構トラック）の中に記憶された全又は完全な数のデータ要素のうちの一部のみを更新するように実行される。しかしながら本発明は、データ構造（フォーマット）ならびに実際のデータ内容の両方を更新すべきであるという表示をホスト又はデータ処理ユニット12が提供した場合といったような或る種のケースにおいては、全てのデータ及びデータフォーマット又は構造が新しい書込み要求で更新されることから、第1のデータ記憶システム14からの旧データの読取りを削除することができる、ということも認識している

。しかしながらこのようなデータ及びフォーマット書込み指令は非常に稀であることから、好ましい実施形態では、各々の読取り要求が書込み要求を第1のデータ記憶システム14から読みとらせるようにする、ということが考慮されている。

本発明の方法は同様に、第2の又は新しいデータ記憶システム16が、ホスト12及び第2のデータ記憶システム16の間の「チャンネル」にサービス提供しているチャンネルプロセスによって引き起こされるデータ転送又は移送とは無関係に

又はこれと並行して第1のデータ記憶システム14と第2のデータ記憶システム16の間の透明な又は「背景」データ移送を提供することをも可能にする。第2の又は新しいデータ記憶システム16を提供することの最終目的は、一般にホスト又はその他のデータ処理システム12に対して強化又は増大した能力を提供することにあるため、第1のデータ記憶システムから第2のデータ記憶システムまでできるかぎり迅速にしかもできるかぎり控え目にデータを移送することが望ましい。

従って、背景の移送又はコピー「タスク」又は「プロセス」27を伴って、本発明に従った第2のデータ記憶システム16内の中央処理ユニットによって実行される一連のソフトウェア命令である本発明の方法（当該技術分野において周知のものであるようなハードウェア及びソフトウェア、例えばEMC Symmetrixシリーズ5500データ記憶システム）は、まず第1に、第2のデータ記憶システム16がホスト又はその他の接続されたデータ処理システム12からのデータ読取り又は書込み要求を処理するのに完全に忙がしいか否かを決定する（ステップ200、図4）。第2のデータ記憶システム16が、ホスト又はデータ処理システム12へ及びこれからこのような要求を取扱うのに完全に忙しいか又は第2のデータ記憶システム16内でその他のデータ入出力（I/O）オペレーションを取扱うのに完全に忙しい場合、さらなる処理は行なわれず、その代わり移送プロセス27は、第2のデータ記憶システム16のオペレーティングシステムから「非繁忙中」又は「使用可能」という表示を待つ。

第2のデータ記憶システム16が、ホスト又はデータ処理装置12からの要求

又は内部の入出力（I/O）要求を取扱うのに忙がしくなくなった時点で、第2のデータ記憶システム16はデータマップ／テーブル24を読取り（ステップ202）、第1のデータ記憶システム14から第2のデータ記憶システム16までコピーされなかったのはどのデータ要素かを決定する（ステップ204）。

前述の通り、第2のデータ記憶システム16の初期構成の間、第2のデータ記憶システムが「オンライン」に来る前に、ユーザー又はシステムエンジニアは、第1のつまり古い方のデータ記憶システム14のデータ記憶システム構成を正確にエミュレートする（すなわち同じメモリーアドレスを有する）ように第2のデ

ータ記憶システム16内のデータ記憶場所17a-17nの少なくとも一部分を構成するためにパーソナルコンピュータ又はその他の入力装置といったシステム構成装置26を利用することになる。一般に、新しいつまり第2のデータ記憶システム16は第1のつまり「旧」データ記憶システム14よりも大きい記憶容量をもつことになり、そのため付加的な記憶領域又は場所が利用可能となる。従って、第1のデータ記憶システム14が、各々一定数のトラック又はレコードをもつ予め定められた数の駆動機構又はボリュームを含んでいる場合、第2のデータ記憶システムは、このような構成を模倣するように構成されることになる。

第2のデータ記憶システム16は、旧のつまり第1のデータ記憶システム14から少なくとも1つのデータ要素（例えばトラック）がコピーされなかったことをひとたび決定したならば、そのデータ要素について第1のデータ記憶システム14に対し要求を発出する（ステップ206）。第2のデータ記憶システム16は、ひとたび受理した時点で、第2のデータ記憶システム16上（標準的にはキャッシュメモリ18内）にデータを記憶し（ステップ208）、第2のデータ記憶システムデータマップ／テーブル24を更新し（ステップ210）、ホスト又はその他のデータ処理システム12からのデータ読取り又は書込み保留要求があるか否かを決定するためにステップ200に戻る。

1つの実施形態においては、本発明は、第1のデータ記憶システム14から第2のデータ記憶システム16までデータを「先取りする」ことが望ましいことである可能性があるということを考慮している。例えば、移送又はコピープロセス

27は、第1のデータ記憶システム14に固有の指令を用いて、第1のデータ記憶システム14に第2のデータ記憶システム16のキャッシュメモリ18に対し一定数のデータ要素を取り出し又は「先取り」し続けさせるべく第1のデータ記憶システム14に対して先取り又は「順次」データアクセス要求又は指令を発出することができる。このような先取りは、データ記憶システム間で移送されなくてはならない「読取り」指令の数を大幅に減少させることによって、第1及び第2のデータ記憶システム14、16の間のデータ転送を著しくスピードアップさせることができる。

もう1つの実施形態においては、移送プロセス27は、ホスト12からの単数又は複数の読取り要求が一連のこのような読取り要求の一部であることを決定することができる。このようなケースにおいては、チャネルプロセス27はホスト12により要求されているデータの現行のアドレスを取り、これを予め定められた数だけ増大させることができる。例えば、ホスト12が現在アドレス「411」からのデータを要求している場合、チャネルプロセス25は、アドレス411にあるデータについて第1のデータ記憶システム14に対し読取り要求を発出することになる。一般に、同時に、チャネルプロセスは、アドレス「413」からデータを先取りするか又は移送し始めるため移送プロセス27に対し表示を渡すことになる。かくして移送プロセス27は、第2のデータ記憶システム16が確実に、第1のデータ記憶システム14からの実際のデータ要求及びチャネルプロセス25を「リード」するようにするために用いられることになる。チャネルプロセス25は、アドレス411及び412にあるデータについてのホスト12からの要求を取扱う。その後続く要求はすでに第2のデータ記憶システム16内でインキャッシュ状態となり、第2のデータ記憶システム16によって迅速に取扱われる。

従って、本発明は、第1のデータ記憶システム上に記憶されたデータへのアクセス上の時間的損失が実質的に全く無い状態で既存のホスト又はその他のデータ処理装置に対して新しいつまり第2のデータ記憶システムを接続できるようにする独特のデータ記憶システム及び方法を提供する。ホスト又はその他の接続され

たデータ処理装置に対するデータの実時間オンライン利用可能性を維持しながら、本発明のシステム及び方法は、現行の及び過去のデータに対するアクセス可能性を提供する。背景オペレーションを用いて、データが第1のつまり古い方のデータ記憶装置から新しいデータ記憶システムへ移送される。

本発明は好ましくは、ソフトウェアの形で実現されるが、当事者であれば本発明をその範囲から逸脱することなくハードウェア又はハードウェアとソフトウェアのさまざまな組合わせの形でも実現できるということが評価できるため、このことが本発明の制限となることはない。

当事者による修正及び置換は、以下のクレームによってのみ制限される本発明の範囲内にあるものとみなされる。

【図 1】

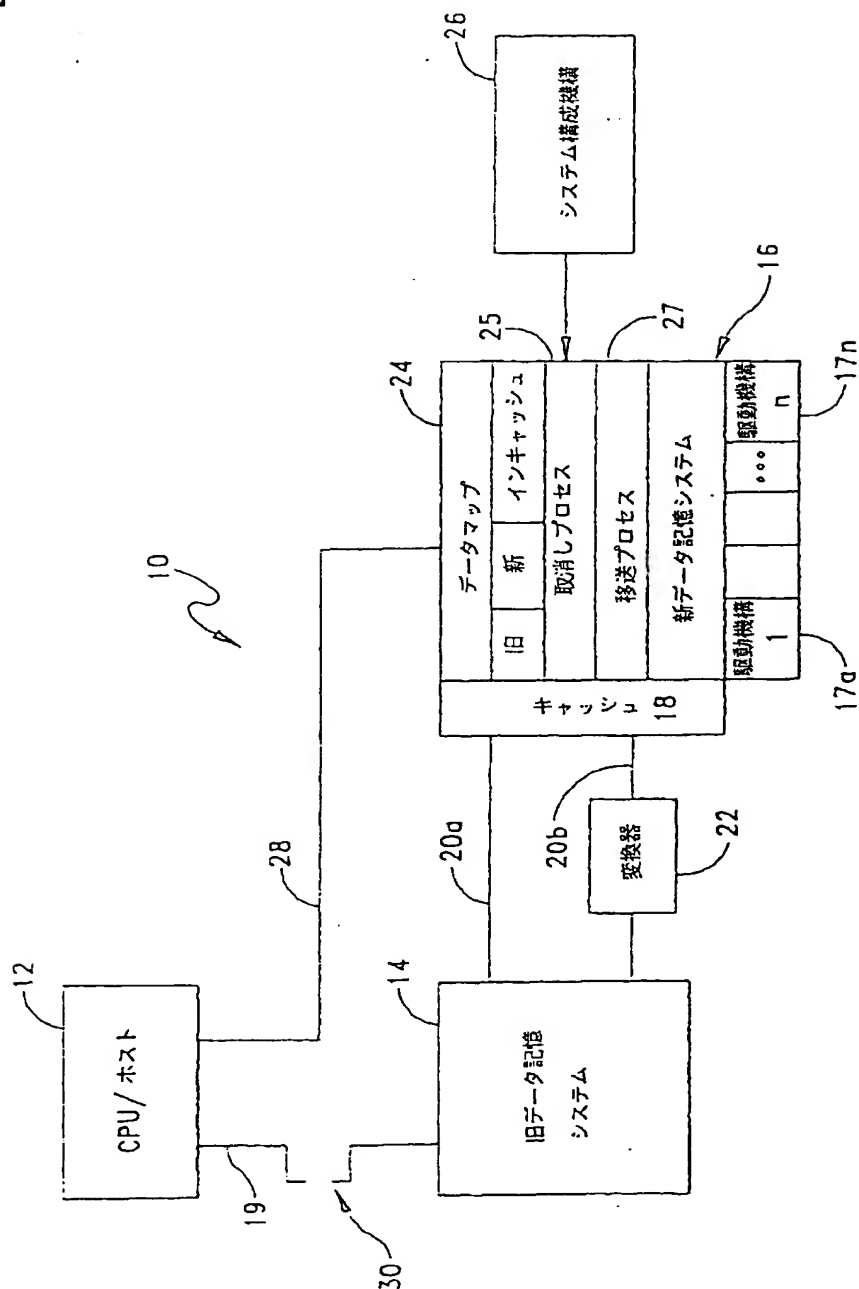


FIG. 1

【図2】

24

58

データマップ/テーブル

インキャッシュ		書き込み保留	その他	新アドレス	移送要
装置 X	X	X		...	はい
ボリューム 1	X	X			はい
トラック 1					はい
トラック 2		X			いいえ
⋮		⋮		⋮	⋮
トラック N					はい
ボリューム 2					はい
⋮		⋮		⋮	⋮

50

52

54a

54b

54c

56

60

64

62

61

FIG. 2

【図3】

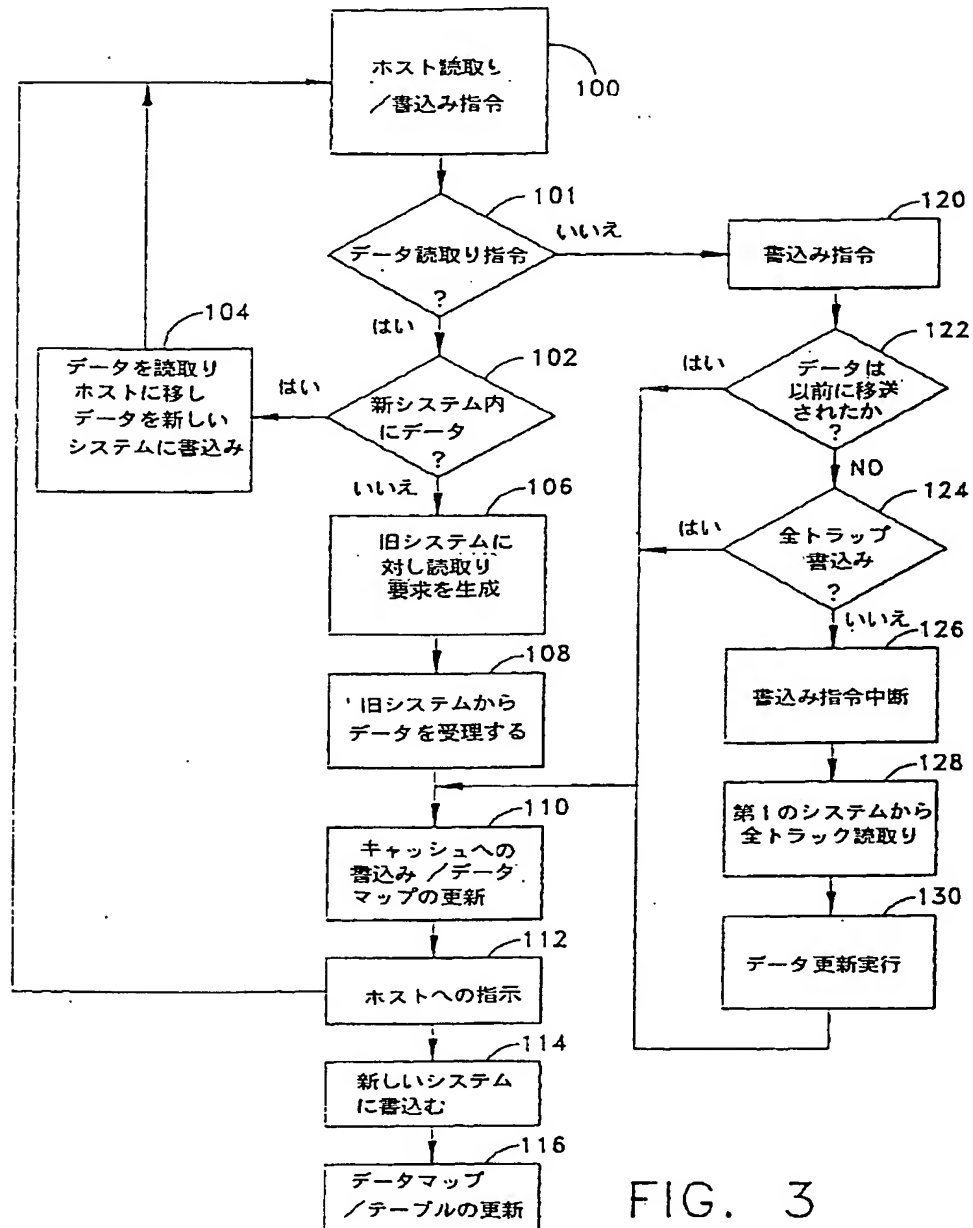


FIG. 3

【図4】

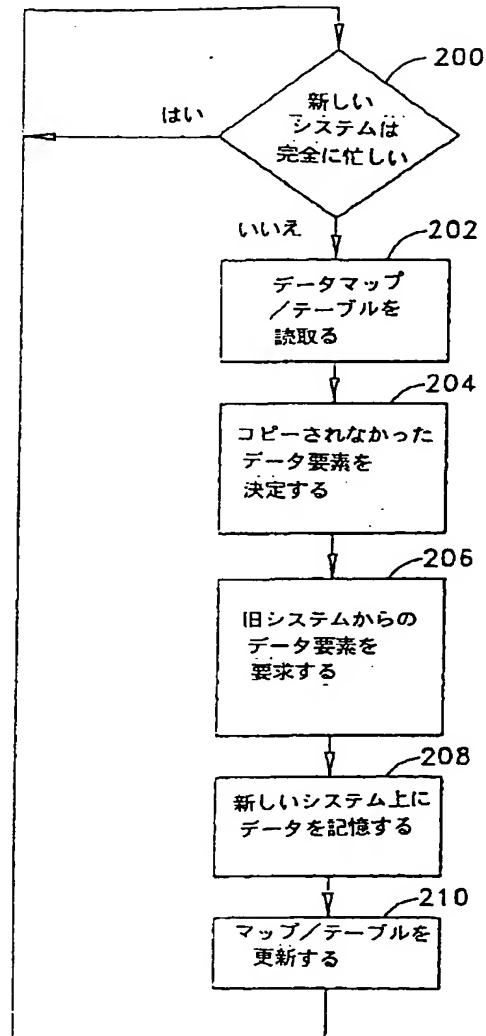


FIG. 4

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/13781

A. CLASSIFICATION OF SUBJECT MATTER IPC(6) : G06F 12/00, 13/00 US CL : 364/232.2, 248.1, 238.4; 395/182.04, 600 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 364/232.2, 248.1, 238.4; 395/182.04, 600 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) APS, Maya, JPOAPS, TDB, EPO, Proquest, Computer Select		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US, A 3,771,137 (BARNER ET AL) 06 November 1973, see col. 2, lines 34-36; col. 2, lines 31-34; col. 2, lines 51-58; col. 2, lines 53-54; col. 4, lines 58-62; col. 9, lines 39-47; col. 2, lines 31-55.	1-33
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family		
Date of the actual completion of the international search 01 NOVEMBER 1996		Date of mailing of the international search report 19 NOV 1996
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230		Authorized officer Thomas C. Lee <i>Jon Hill</i> Telephone No. (703) 305-9717

Form PCT/ISA/210 (second sheet) (July 1992)*